

Nonlinear Filtering for Stochastic Volatility Models with Heavy Tails and Leverage

A.E. Clements, S.I. White.

School of Economics and Finance, Queensland University of Technology

September 6, 2004

Abstract

This paper develops a computationally efficient filtering based procedure for the estimation of the heavy tailed SV model with leverage. While there are many accepted techniques for the estimation of standard SV models, incorporating these effects into an SV framework is difficult. Simulation evidence provided in this paper indicates that the proposed procedure outperforms competing approaches in terms of the accuracy of parameter estimation. In an empirical setting, it is shown how the individual effects of heavy tails and leverage can be isolated using standard likelihood ratio tests.

1 Introduction

Much research attention has been directed at issues surrounding the estimation of the log-normal stochastic volatility (SV) model of Taylor (1986). While this model is theoretically appealing, there are many practical issues surrounding the estimation of its parameters. In recent times, a number of extensions to the standard SV model have been proposed, reflecting well known empirical features of financial time series. The SV model has been extended to incorporate, either individually or jointly, heavy tailed error distributions and the leverage effect.

In its most basic form the SV model assumes that return and variance processes are driven by uncorrelated Gaussian innovations. To allow for heavy-tailed return distributions, Liesenfeld and Jung (2000) allow the return process to be based on either t or *ged* distributed errors. The leverage effect is incorporated by Harvey and Shephard (1996) who allow for correlation between return and variance innovations by conditioning the prediction of log-variance on the sign of the return in the QML estimator. Finally, Jacquier, Polson, and Rossi (2004) jointly allow for the leverage effect and t distributed return innovations with their heavy tailed SV model with leverage.

Although models have been proposed which allow for both leverage and heavy tails, numerous issues still remain unresolved. It is not obvious how the Simulated Maximum Likelihood (SML) procedure used by Liesenfeld and Jung (2000) to estimate the heavy tailed SV model can easily be extended to incorporate the leverage effect. Conversely, the QML procedure used by Harvey and Shephard (1996) to allow for leverage is not amenable to heavy tailed return distributions due to its rigid structural form. Furthermore, whilst the Monte-Carlo Markov Chain (MCMC) procedure of Jacquier et al (2004) is very flexible, it is very computationally intensive and simulation reveal that the procedure produce biased estimates of the leverage parameter.

To address these issues, this paper proposes a non-linear filtering procedure

for the maximum likelihood estimation of a heavy tailed SV model with leverage. The proposed approach utilises the computationally efficient discretised non-linear filtering procedure (DNF) of Clements, Hurn and White (2004) to directly approximate the non-linear filtering problem arising due to correlation between return and volatility innovations. The necessary filtering equations are extensions of those proposed in Kitigawa (1987).

This paper is structured as follows. Section 2, introduces the heavy tailed SV model with leverage and its associated state space representation. Section 3 outlines the specific filtering equations necessary in estimating the heavy tailed SV model with leverage. Furthermore, it is shown how these filtering equations can be implemented using the DNF. Section 4 reports simulation results for the DNF procedure applied to the heavy tailed SV model with leverage, allowing for comparisons to alternative estimation procedures such as the MCMC procedure of Jacquier et al (2004). Additionally, Section 5 applies this estimation procedure to equity index and currency returns. Section 6 provides concluding remarks.

2 The heavy tailed SV model with leverage

In the standard lognormal SV model, returns $\{y_t\}_{t=1}^T$ are generated by,

$$\begin{aligned} y_t &= \exp\left(\frac{x_t}{2}\right) u_t \\ x_t &= \alpha + \beta x_{t-1} + \sigma_v w_t \end{aligned} \tag{1}$$

where $x_t = \log(\sigma_t^2)$, yielding $\exp\left(\frac{x_t}{2}\right)$ as the time t conditional standard deviation of y_t . In its simplest form, u_t and w_t are treated as two uncorrelated white noise processes implying independence between return and volatility innovations. In this state-space form, the expression for y_t is the observation equation with the state equation governing the behaviour of x_t .

However, in the context of equity returns, Black (1976) and Campbell and Hentschel (1992) theoretically justify the presence of negative correlation between u_t and w_t . This empirical pattern is known as the leverage effect and

reflects the fact that as rates of return fall, volatility (risk) rises. Incorporating such effects into GARCH style models is relatively straightforward given that volatility is treated as a deterministic function of historical information (such as the sign of past returns). Commonly used asymmetric GARCH style models include EGARCH and GJR models of Nelson (1991) and Glosten, Jagannathan and Runkle (1993) respectively. For a summary of a wider array of related models, see Hentschel (1995) and Pagan (1996). Dealing with the leverage effect in an SV context is somewhat more difficult given the given that conditional volatility is an unobserved stochastic variable within the state-space form outlined in equation 1.

To allow for a leverage effect in the SV framework, the dynamics of equation 1 must be generalized to allow $E[u_t, w_t] = \rho$. By permitting $E[u_t, w_t] = \rho$ the likelihood of y_t involves the joint distribution of y_t and x_t . To evaluate the likelihood function of the general SV model of equation 1, the marginal distribution of y_t conditional on x_t is required. The marginal distribution of y_t can be obtained by making the substitution:

$$u_t = \rho w_t + \sqrt{1 - \rho^2} u_t^* \quad (2)$$

where $E[u_t^*, w_t] = 0$. Substituting equation 2, into equation 1 we obtain the following form for the return equation:

$$y_t = \exp\left(\frac{x_t}{2}\right) \left(\rho w_t + \sqrt{1 - \rho^2} u_t^*\right). \quad (3)$$

From equation 1, w_t can be rewritten as:

$$w_t = \frac{x_t - \alpha - \beta x_{t-1}}{\sigma_v}. \quad (4)$$

Substituting equation 4 into equation 3 eliminates the random term w_t leading to the following observations and state equations,

$$\begin{aligned} y_t &= \frac{\exp\left(\frac{x_t}{2}\right) \rho (x_t - \alpha - \beta x_{t-1})}{\sigma_v} + \exp\left(\frac{x_t}{2}\right) \sqrt{1 - \rho^2} u_t^* \\ x_t &= \alpha + \beta x_{t-1} + \sigma_v w_t \end{aligned} \quad (5)$$

Note that when $\rho = 0$, equation 5 collapses to equation 1. Under the assumption that $u_t \sim N(0, 1)$, equation 5 becomes the standard SV model with leverage where the marginal distribution of y_t is:

$$p(y_t | x_t, x_{t-1}, y_{t-1}) = \frac{1}{\sqrt{2\pi \exp(x_t)(1 - \rho^2)}} \exp\left(-\frac{(y_t - \mu_t)^2}{2 \exp(x_t)(1 - \rho^2)}\right) \quad (6)$$

and

$$\mu_t = \frac{\exp\left(\frac{x_t}{2}\right) \rho (x_t - \alpha - \beta x_{t-1})}{\sigma_v} \quad (7)$$

One method for allowing for y to be generated from a heavy tailed error distribution is to allow $u_t \sim t_v(0, 1)$ (where v is the degrees of freedom). In this case we get the heavy tailed SV model with leverage similar to Jacquier et al (2004), but here the marginal distribution of y is:

$$p(y_t | x_t, x_{t-1}, y_{t-1}) = [\pi(v-2)(1 - \rho^2) \exp(x_t)]^{-\frac{1}{2}} \frac{\Gamma((v-1)/2)}{\Gamma(v/2)} \left[1 + \frac{(y_t - \mu_t)^2}{\exp(x_t)(1 - \rho^2)(v-2)}\right]^{-\frac{v+1}{2}} \quad (8)$$

where μ_t is defined in equation 7. Under the constraint that $\rho = 0$, equations 6 and 8 represent the standard SV and heavy tailed SV models respectively.

Given that such models described above are latent variable processes, even in the simple case of equation 1 the true likelihood function is analytically intractable. The following section will outline how the likelihood of $\{y_t\}_{t=1}^T$ can be approximated by extending the general filtering framework proposed by Kitigawa (1987) to deal with the new state space form of equation 5.

3 Nonlinear filtering for SV models

To estimate the standard SV model (where $\rho = 0$) of equation 1 in a nonlinear filtering framework, two definitions are required,

$$y_t \sim r(\cdot | x_t, y_{t-1}), \quad x_t \sim q(\cdot | x_{t-1}, y_{t-1}) \quad (9)$$

where $r(\cdot)$ and $q(\cdot)$ are derived from the observation and state equations. Kitigawa (1987) outlines the relevant filtering equation necessary to evaluating the likelihood of $\{y_t\}_{t=1}^T$. Given a parameter vector θ , the resulting prediction-update filter is defined in the following 2 steps.

Prediction step

The one-step ahead prediction of the distribution of x_t conditional on y_{t-1} , $f(x_t | y_{t-1}, \theta)$, is given by

$$f(x_t | y_{t-1}, \theta) = \int_{-\infty}^{\infty} q(x_t | x_{t-1}, \theta) f(x_{t-1} | y_{t-1}, \theta) dx_{t-1}. \quad (10)$$

Update step

The probability distribution of the state variable at time t , conditional on information at time t is then given by

$$f(x_t | y_t, \theta) = \frac{r(y_t | x_t, y_{t-1}, \theta) f(x_t | y_{t-1}, \theta)}{f(y_t | y_{t-1}, \theta)}. \quad (11)$$

where the denominator of equation (11) is the likelihood of observing y_t given y_{t-1} and θ

$$f(y_t | y_{t-1}, \theta) = \int_{-\infty}^{\infty} r(y_t | x_t, \theta) f(x_t | y_{t-1}, \theta) dx_t.$$

When the standard SV model is generalised such that $\rho \neq 0$, the marginal distribution of y_t now depends not only on x_t but also on x_{t-1} . Hence, the state space representation of equation 9 becomes:

$$y_t \sim r(\cdot | x_t, x_{t-1}, y_{t-1}), \quad x_t \sim q(\cdot | x_{t-1}, y_{t-1}). \quad (12)$$

To reflect this more general state space form, the prediction and update equations must be modified as follows.

Prediction Step:

Since marginal distribution of y_t depends jointly on x_t and x_{t-1} , the predicted joint distribution of x_t and x_{t-1} must be evaluated¹. The prediction step therefore becomes:

$$p(x_t, x_{t-1} | y_{t-1}, \theta) = q(x_t | x_{t-1}, y_{t-1}, \theta) p(x_{t-1} | y_{t-1}, \theta) \quad (13)$$

Update Step:

The form of the probability distribution of the x_t , conditional on information at time t is then given by

$$\begin{aligned} p(x_t | y_t, \theta) &= \int_{-\infty}^{\infty} p(x_t, x_{t-1} | y_t, \theta) \partial x_{t-1} \\ &= \frac{\int_{-\infty}^{\infty} p(x_t, x_{t-1}, y_t | y_{t-1}, \theta) \partial x_{t-1}}{p(y_t | y_{t-1}, \theta)} \\ &= \frac{\int_{-\infty}^{\infty} r(y_t | x_t, x_{t-1}, y_{t-1}, \theta) p(x_t, x_{t-1} | y_{t-1}, \theta) \partial x_{t-1}}{p(y_t | y_{t-1}, \theta)} \end{aligned} \quad (14)$$

Likelihood

The likelihood of y_t is now represented by:

$$\begin{aligned} p(y_t | y_{t-1}, \theta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y_t, x_t, x_{t-1} | y_{t-1}, \theta) \partial x_t \partial x_{t-1} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(y_t | x_t, x_{t-1}, y_{t-1}, \theta) p(x_t, x_{t-1} | y_{t-1}, \theta) \partial x_t \partial x_{t-1} \end{aligned} \quad (15)$$

¹Although not directly required for this filtering procedure, it is possible to generate a one-step ahead prediction of $p(x_t | y_{t-1})$ by integrating $p(x_t, x_{t-1} | y_{t-1})$ with respect to x_{t-1} .

The ability to evaluate the likelihood of $\{y_t\}_{t=1}^T$ rests on the ability of an algorithm to generate estimates of $p(x_t | y_t, \theta)$ and $p(y_t | y_{t-1}, \theta)$ in equations 14 and 15. In the standard SV case of equation 1, only single integrals need to be evaluated. While Kitigawa (1987) and Fridman and Harris (1998) propose integration schemes to deal with the standard SV case of equations (10 and 11), in the current context equations (14 and 15) require double integrals to be evaluated. While the integration approaches proposed by Kitigawa (1987) and Fridman and Harris (1998) could be extended to deal with such double integrals, the associated computational burden would render estimation impractical. To provide a computationally tractable algorithm for evaluating the likelihood function, an extension to the discretised non-linear filtering (DNF) procedure of Clements *et al.* (2004) is proposed. The utility of the DNF in this context is that the likelihood function may be evaluated without the use of directly numerically evaluating the double integrals in equations 14 and 15.

3.1 Discretised Non-Linear Filtering

The Discrete Non-Linear Filter (DNF) proposed by Clements *et al.* (2004) solves the non-linear filtering equations based on a discretisation of state-space. This allows the likelihood function of a continuously valued latent variable process to be evaluated in a similar manner to Markov models for discrete valued time series, see MacDonald and Zucchini (1997). In doing so, this avoids the use of numerical integration schemes such as those proposed by Kitagawa (1987) and Fridman and Harris (1998).

Under the DNF approach, the *pdf* of the latent variable, x , is approximated by computing the probability of observing x within a set of discrete intervals (a histogram) as opposed to the linear spline approach suggested by Kitagawa (1987). In discretising state-space, N adjacent intervals in x space are defined,

bounded by $w^1 \dots w^{N+1}$, and centered on the points $x^1 \dots x^N$ where

$$x^i = \frac{w^i + w^{i+1}}{2}. \quad (16)$$

In general terms, the probability of observing x within the interval centered on x^i , i.e. $x \in (w^i, w^{i+1}]$ is given by

$$p(x \in (w^i, w^{i+1}]) = \int_{w^i}^{w^{i+1}} f(x) dx \approx p(x^i) \quad (17)$$

where $f(x)$ is the continuous probability distribution of the of the unobserved state variable x . The series of $p(x^i)$, $i = 1, \dots, N$ represent a discretised approximation to the continuous distribution $f(x)$. Based on this discrete approximation, the DNF captures the evolution of the state variable through time given definitions of a time-invariant set of transition probabilities and a set of conditional likelihoods.

Transitional probabilities

Following from equation 12 the transitional density of x may be discretised into a set of transitional probabilities. Given that the state space is defined over N adjacent intervals it is possible to compute an $N \times N$ matrix of time-invariant transition probabilities, \hat{q} . The elements of this matrix, $\hat{q}^{i,j} \forall i, j = 1, \dots, N$, represent the probability of x migrating from the interval centred on x^j at time $t - 1$, to the interval centred on x^i at time t and is given by

$$\hat{q}^{i,j} = \delta q(x_t^i | x_{t-1}^j, \theta) \quad (18)$$

where δ is the interval width. In the case where $q(\cdot)$ is a normal distribution,

$$\hat{q}^{i,j} = \frac{\delta}{\sqrt{2\pi\sigma_v^2}} \exp\left(-\frac{(x^i - \alpha - \beta x^j)^2}{2\sigma_v^2}\right).$$

Conditional likelihoods

Following from equation 12, the likelihood of observing y_t conditional on x being within each discrete interval is found. The $T \times N$ likelihood matrix containing elements, $\hat{r}_t^i \forall i = 1, \dots, N$, is defined by

$$\hat{r}_t^{i,j} = r(y_t | x_t^i, x_{t-1}^j, \theta) \quad (19)$$

In the standard SV model with leverage ($u_t \sim N(0, 1)$), $\hat{r}_t^{i,j}$ is given by

$$\hat{r}_t^{i,j} = \frac{1}{\sqrt{2\pi \exp(x^i)(1 - \rho^2)}} \exp\left(-\frac{(y_t - \mu^{i,j})^2}{2 \exp(x^i)(1 - \rho^2)}\right) \quad (20)$$

where

$$\mu^{i,j} = \frac{\exp(\frac{x^i}{2})\rho(x^i - \alpha - \beta x^j)}{\sigma_v} \quad (21)$$

In the heavy tailed SV model with leverage ($u_t \sim t_v(0, 1)$), $\hat{r}_t^{i,j}$ is given by

$$\hat{r}_t^{i,j} = [\pi(v-2)(1-\rho^2)\exp(x^i)]^{-\frac{1}{2}} \frac{\Gamma((v-1)/2)}{\Gamma(v/2)} \left[1 + \frac{(y_t - \mu_t^{i,j})^2}{\exp(x^i)(1-\rho^2)(v-2)}\right]^{-\frac{v+1}{2}} \quad (22)$$

Based on this set of conditional likelihoods and the time-invariant matrix of transition probabilities, the DNF proceeds with the following steps.

Prediction Step

The predicted joint probability of observing $x \in (w^i, w^{i+1}]$ at time t and $x \in (w^j, w^{j+1}]$ at time $t-1$ is given by:

$$\begin{aligned} P_t^{i,j} &= p(x_t^i, x_{t-1}^j | y_{t-1}, \theta) \\ &= q(x_t^i | x_{t-1}^j, y_{t-1}, \theta) p(x_{t-1}^j | y_{t-1}, \theta) \\ &= \hat{q}^{i,j} \cdot U_{t-1}^j. \end{aligned} \quad (23)$$

Update Step

The updated probability of observing $x \in (w^i, w^{i+1}]$ at time t , is defined as

$$\begin{aligned}
U_t^i &= p(x_t^i | y_t, \theta) \\
&= \frac{\sum_{j=1}^N r(y_t | x_t^i, x_{t-1}^j, y_{t-1}, \theta) p(x_t^i, x_{t-1}^j | y_{t-1}, \theta)}{p(y_t | y_{t-1}, \theta)} \\
&= \frac{\sum_{j=1}^N \hat{r}_t^{i,j} \cdot P_t^{i,j}}{p(y_t | y_{t-1}, \theta)}
\end{aligned} \tag{24}$$

Likelihood

The denominator of equation (24) is the likelihood of observing y_t , given by

$$\begin{aligned}
p(y_t | y_{t-1}, \theta) &= \sum_{i=1}^N \sum_{j=1}^N r(y_t | x_t^i, x_{t-1}^j, y_{t-1}, \theta) p(x_t^i, x_{t-1}^j | y_{t-1}, \theta) \\
&= \sum_{i=1}^N \sum_{j=1}^N \hat{r}_t^{i,j} \cdot P_t^{i,j}
\end{aligned} \tag{25}$$

The advantage of the DNF procedure in this context is that the numerical integration required to evaluate equations 14 and 15 have been replaced by matrix operations used in a discrete valued Markov setting. For the DNF to be initialised, the prediction of the state probabilities at time $t = 0$ need to be selected. The state probabilities are initialised by discretising the unconditional distribution of the state variable such that

$$U_0^i = \int_{w^i}^{w^{i+1}} f(x | \theta) dx. \tag{26}$$

Smoothing

To generate estimates of the latent volatility process at any time given the entire data series, fixed interval smoothing can be undertaken. The smoothing algorithm in the presence of leverage and heavy tails is the same as the smoothing algorithm for the uncorrelated case (see Clements *et al.* (2004)). To implement the smoothing algorithm the one-step ahead prediction of the distribution of x_t , $p(x_t | y_{t-1}, \theta)$ needs to be evaluated. The prediction step in equation

23 however represents the prediction of the joint distribution of x_t and x_{t-1} , $p(x_t, x_{t-1} | y_{t-1}, \theta)$. The one-step ahead prediction given in terms of $p(x_t^i | y_{t-1}, \theta)$ can be found as follows:

$$p(x_t^i | y_{t-1}, \theta) = \sum_{j=1}^N P_t^{i,j} \quad (27)$$

where $P_t^{i,j}$ is given in equation 23.

4 Simulation Evidence

This section contains the results of two simulation studies which examine the performance of the DNF approach with respect to estimating SV models with leverage. The first simulation study considers the performance of the standard SV model with leverage whilst the second considers the heavy-tailed SV model with leverage.

4.1 Standard SV model with leverage

Simulation results contained in Table 1 assess the ability of the DNF to estimate the parameters of the standard SV model with leverage. This study moves beyond the one parameter set considered in Jacquier *et al.* (1994) to include different levels of correlation between return and log-variance innovations. The parameter set used in this study is $\{\alpha, \beta, \sigma_v\} = \{-0.363, 0.95, 0.26\}$ with $\rho = \{0, -0.3, -0.5, -0.7\}$. Parameters are estimated on 500 simulated samples of size $T = 500, 1000$ and 2000 . The number of intervals used in the discretisation of the log-variance state space is $N = 25$, with the state space uniformly discretised over $\frac{\alpha}{(1-\beta)} \pm 3 \frac{\sigma_v}{\sqrt{(1-\beta^2)}}$. All estimations are performed with the following parameter constraints: $\alpha \in (-100, 100)$, $b \in (.2, .99)$, $\sigma_v \in (.01, 1)$, $\rho \in (-.99, .99)$. Additionally, the unconditional level of x_t is constrained to be representative of the unconditional return variance, h , such that $\frac{\alpha}{(1-\beta)} \in (\log(h/3), \log(3h))$.

Examination of the simulation results presented in Table 1 reveals that as ρ

$T = 500$	α	β	σ_v	ρ
	-.363	0.95	0.26	
$\rho = 0$	-0.395 (0.221)	0.946 (0.030)	0.257 (0.065)	-0.0023 (0.166)
$\rho = -0.3$	-0.386 (0.198)	0.948 (0.027)	0.258 (0.057)	-0.325 (0.149)
$\rho = -0.5$	-0.398 (0.189)	0.946 (0.026)	0.268 (0.057)	-0.517 (0.132)
$\rho = -0.7$	-0.367 (0.123)	0.950 (0.017)	0.261 (0.046)	-0.722 (0.103)

$T = 1000$	α	β	σ_v	ρ
	-.363	0.95	0.26	
$\rho = 0$	-0.339 (0.119)	0.954 (0.016)	0.258 (0.043)	-0.0009 (0.114)
$\rho = -0.3$	-0.349 (0.125)	0.953 (0.017)	0.261 (0.041)	-0.305 (0.109)
$\rho = -0.5$	-0.351 (0.106)	0.952 (0.014)	0.260 (0.038)	-0.525 (0.093)
$\rho = -0.7$	-0.342 (0.087)	0.953 (0.012)	0.262 (0.031)	-0.716 (0.072)

$T = 2000$	α	β	σ_v	ρ
	-.363	0.95	0.26	
$\rho = 0$	-0.314 (0.098)	0.957 (0.013)	0.255 (0.031)	0.001 (0.074)
$\rho = -0.3$	-.320 (0.092)	0.956 (0.012)	0.258 (0.030)	-0.309 (0.073)
$\rho = -0.5$	-0.329 (0.082)	0.955 (0.011)	0.259 (0.027)	-0.517 (0.0645)
$\rho = -0.7$	-0.332 (0.070)	0.955 (0.010)	0.262 (0.024)	-0.713 (0.049)

Table 1: Simulation Results for the DNF applied to the SV models with leverage. Mean parameter estimates are reported along RMSE in brackets

tends away from zero all parameter estimates become more precise². While there is no directly comparable simulation evidence regarding the estimation of the leverage parameter, Harvey and Shephard (1996) report simulation results for a similar experiment using QML with sample sizes of 500 and 1000 and parameter set $\{\alpha, \beta, \sigma_v\} = \{0, 0.975, 0.1\}$. At a sample size of 500, $RMSE = 0.497$ and 0.465 for $\rho = 0$ and $\rho = -0.3$ respectively. For a sample size of 1000, $RMSE = 0.325$ and 0.298 for corresponding values of ρ . Whilst these results are not directly comparable to those in Table 1, the DNF procedure exhibits approximately one third of the $RMSE$ of the QML procedure for the same values of ρ .

4.2 Heavy tailed SV model with leverage.

The simulation study of Jacquier et al (2004) considers the estimating the heavy tailed SV model with leverage for $\{\alpha, \beta, \sigma_v, \nu, \rho\} = \{-363, 0.95, 0.26, 10, -0.6\}$. In this study, 500 series of length 1000 are simulated from the heavy tailed SV model with leverage. This experiment is replicated using the DNF procedure with the mean and $RMSE$ of the resulting parameter estimates compared with those reported in Jacquier *et al.* (2004).

For the purposes of parameter estimation the same parameter constraints as in section 4 are used, along with the additional constraint of $\nu \in (2.1, 200)$. Setting $\min(\nu) = 2.1$ ensures that for all likelihood evaluations the variance exists, furthermore, setting $\max(\nu) = 200$, leads to u_t being approximately normally distributed and ensures that the evaluation of the $t - pdf$ is numerically feasible.

It is seen from Table 2 that the DNF procedure estimates both ν and ρ with less bias than the MCMC procedure. Most strikingly, bias in the estimation of the ρ parameter falls from 0.18 for the MCMC approach to 0.025 for the DNF procedure. Furthermore, the DNF provides dramatic reduction in $RMSE$, a result that is dominated by the bias reduction. Here, $RMSE$ for the ν parameter is not reported, as Jacquier et al (2004) only report the average of the 1st and 3rd

²This finding is consistent with Harvey and Shephard (1996).

<i>True Value</i>	β (0.95)	σ_v (0.26)	ν (10)	ρ (-0.6)
<i>MCMC</i>	0.94 (0.025)	0.27 (0.039)	14	-0.42 (0.19)
<i>DNF</i>	0.952 (0.014)	0.264 (0.042)	10.387	-0.5754 (0.083)

Table 2: Simulation results for the DNF procedure applied to the heavy tailed SV model. Mean parameter estimates are reported along with RMSE in brackets.

quartiles of the posterior distribution which is not comparable to *RMSE*. Even though the *RMSE* is not reported, the parameter v is estimated with less bias using the DNF procedure.

5 Empirical Example

While the leverage effect was initially motivated from the perspective of equity returns, the importance of both leverage and heavy-tail features in an SV context will be considered with respect to equity index and spot currency returns. Equity index returns are represented by daily returns on the S&P500 Composite Index, from 5 September 1996 to 16 August 2003 (2000 daily observations). Currency returns considered, are daily returns on the Japanese Yen, US Dollar (JPY/USD) spot exchange rate from 29 November 1996 to 30 July 2004 (2000 daily observations).³

To evaluate the individual and joint effects of allowing for leverage and heavy tails the heavy tailed SV model with leverage is estimated with a number of parameter restrictions. For the standard SV model $\rho = 0$ and $v = \infty$, the heavy tailed SV model $\rho = 0$ and the SV model with leverage $v = \infty$. Table 3 outlines the parameter estimates for the four SV models nested within the heavy tailed SV with leverage model for both the S&P 500 and JPY/USD series respectively. This empirical study utilises the parameter constraints discussed in Section 4.

³Returns for both series are expressed in percentage terms. Furthermore, each is a zero mean series.

<i>S&P 500</i>				
	<i>SV</i>	<i>SV - t</i>	<i>SV - ρ</i>	<i>SV - t&ρ</i>
α	$4.01E - 3$ ($7.8E - 4$)	$3.91E - 3$ ($4.0E - 3$)	$8.00E - 3$ ($5.0E - 4$)	$7.99E - 3$ ($5.7E - 4$)
β	0.975 ($1.1e - 3$)	0.977 ($4.2E - 3$)	0.982 ($7.5E - 4$)	0.982 ($7.8E - 4$)
σ_v	0.155 ($2.9E - 3$)	0.146 (0.018)	0.143 ($2.3E - 3$)	0.143 ($2.2E - 3$)
ρ	0	0	-0.786 ($6.4E - 3$)	-0.786 ($5.7E - 3$)
v	∞	33.6 (37.9)	∞	200* (30.7)
<i>Loglike</i>	-3108.5	-3107.9	-3067.8	-3067.8

<i>JPY / USD</i>				
	<i>SV</i>	<i>SV - t</i>	<i>SV - ρ</i>	<i>SV - t&ρ</i>
α	-0.028 ($2.7E - 3$)	-0.014 ($4.6E - 3$)	-0.022 ($2.7E - 3$)	-0.008 ($1.5E - 3$)
β	0.966 ($3.1E - 3$)	0.981 ($3.7E - 3$)	0.971 ($3.2E - 3$)	0.987 ($2.1E - 3$)
σ_v	0.178 ($7.0E - 3$)	0.117 (0.018)	0.152 ($7.6E - 3$)	0.092 ($4.3E - 3$)
ρ	0	0	-0.312 (0.023)	-0.268 (0.042)
v	∞	7.96 (1.63)	∞	7.41 (0.357)
<i>Loglike</i>	-2118.9	-2101.2	-2110.6	-2099.8

Table 3: Parameter estimates for the heavy tailed SV model. An asterisk represents a parameter estimated on the constraint. Bollerslev and Woldridge standard errors are in brackets.

Results in Table 3 for the standard SV model applied to the S&P 500 returns reveal the familiar pattern of high persistence in conditional volatility. Augmenting the standard SV model to include t distributed errors ($SV - t$) in the context of the S&P 500 dataset is not important. In this case there is no significant improvement in the likelihood ($LR = 1.2$, $\chi^2_{1,0.05} = 3.84$). This is reflected in the large estimate of the degrees of freedom for the t -distribution of 33.6. However, it is seen that a leverage effect is clearly an important feature of this dataset. In comparison to the standard SV model, the SV with leverage model ($SV - \rho$) leads to a significant improvement in likelihood ($LR = 81.4$). This pattern is reflected in the significant estimate of $\rho = -0.786$, a common finding when dealing with equity returns. When both leverage and heavy tail effects are considered ($SV - t \& \rho$) there is no significant improvement beyond the $SV - \rho$ model.

Results for the JPY/USD series represent a contrasting example to the S&P 500 returns. In this case, the $SV - t$ leads to a significant increase in the likelihood above that of the standard SV models ($LR = 35.4$, $\chi^2_{1,0.05} = 3.84$). The resulting error distribution with 7.96 degrees of freedom is clearly heavy tailed in comparison to the normal distribution. This is a common result when dealing with currency returns, see Jacquier et al (2004) and Liesenfeld and Jung (2000). When leverage is considered, the $SV - \rho$ model also provides a significant increase in the likelihood relative to the SV model ($LR = 16.6$) associated with a significant estimate of $\rho = -0.312$. While the, $SV - t \& \rho$ model represents an improvement over the $SV - \rho$ model ($LR = 21.6$) it does not provide a significant increase in likelihood over the $SV - t$ model ($LR = 2.8$). These results indicate that heavy-tails are the dominant feature in this JPY/USD series.

6 Concluding remarks

This paper has considered the estimation of a heavy tailed SV model with leverage. While much attention has been paid to the estimation of various SV models,

extensions to include these well known empirical feature of asset returns has been problematic. A non-linear filtering framework to directly evaluate the likelihood of extended SV models has been proposed. To estimate such models in a maximum likelihood setting, the DNF procedure has been tailored to suit this specific problem.

Simulation results suggest that as compared to either MCMC or QML approaches, the DNF estimation procedure more accurately captures both leverage and heavy tail effects. Given that the heavy tail SV model with leverage nests a number of specification, it provides a useful testing framework to isolate the important features of financial asset returns. Empirical results presented in this paper indicate that it is the leverage effect that is most dominant when dealing with a sample of equity index returns, while it is heavy tail features that are dominant when dealing with currency returns.

References

- Black, F., 1976, "Studies in stock volatility changes", Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, 177-181.
- Campbell, J.Y. and Hentschel, L., 1992, "No news is good news: An asymmetric model of changing volatility in stock returns", *Journal of Financial Economics*, 31, 281-318.
- Clements A.E., Hurn A.S. and White S.I. (2004). "Discretised Non-Linear Filtering of Dynamic Latent Variable Models with Application to Stochastic Volatility." Discussion Paper No 179, School of Economics and Finance, Queensland University of Technology.

- Fridman, M. and Harris, L., 1998, "A maximum likelihood approach for non-gaussian stochastic volatility models", *Journal of Business and Economic Statistics*, 16, 284-291
- Glosten, L., Jagannathan, R. and Runkle, D., 1993, "On the relation between the expected value and the volatility of the nominal excess return on stocks" *The Journal of Finance*, 48, 1779-1801.
- Harvey, A.C., Ruiz, E. and Shephard, N., 1994, "Multivariate stochastic variance models", *Review of Economic Studies*, 61, 247-264.
- Harvey, A. and Shephard, N. 1996 "Estimation of an asymmetric stochastic volatility model for asset returns" *Journal of Business and Economic Statistics*, 14, 429-434.
- Hentschel, L, 1995, "All in the Family: Nesting Symmetric and Asymmetric Garch Models", *Journal of Financial Economics*, 39, 71-104.
- Jacquier, E. and Polson, G. and Rossi, P.E., 1994, "Bayesian analysis of stochastic volatility models", *Journal of Business and Economic Statistics*, 12, 371-389.
- Jacquire, E., Polson, N. and Rossi, P., 2004, "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors", *Journal of Econometrics*, 122, 185-212.
- Kitagawa, G., 1987, "Non-gaussian state-space modeling of non-stationary time series", *Journal of the American Statistical Association*, 82, 1032-1062.
- Liesenfeld, R and Jung, R.C., 2000. "Stochastic volatility models: conditional normality versus heavy-tailed distributions." *Journal of applied Econometrics*. 15, 137-160.

Martin, D.R. and Raferty, A.E. (1987). “Robustness, computation and non-Euclidean models” *Journal of the American Statistical Association*, 82, 1044-1050.

MacDonald, I. and Zucchini, W. 1997 *Hidden markov and other models for discrete-valued time series*, Chapman and Hall, London.

Nelson, D., 1991 “Conditional heteroscedasticity in asset returns: A new approach” *Econometrica*, 59, 347-370.

Pagan, A.R., 1996, “The econometrics of financial markets”, *Journal of Empirical Finance*, 3, 1, 15-102.

Ruiz, E., 1994, “Quasi-maximum likelihood estimation of stochastic volatility models”, *Journal of Econometrics*, 63, 289-306.

Taylor S, 1986, *Modelling Financial Time Series*, Wiley, Chichester.